CLAIMS

What is claimed is:

5   1.   A method for using a phased learning approach for determining the closest content serving sites to a fixed location in a computer network including:

in a setup phase:

storing data as to the transit times of each of the content serving sites in a table;

10   in an execution phase:

receiving a DNS request from said fixed location;

determining a response time for each of n fastest content serving sites and m other content serving sites, said n fastest content serving sites chosen by using said data and choosing the n content serving sites with the lowest transit times, said m other content serving sites chosen by selecting new content serving sites as well as randomly selecting old content serving sites;

sending queries to each of the n fastest content serving sites and m other content serving sites, timing said queries so that they arrive at each of the n fastest content serving sites and m other content serving sites at the same time by using said response time for each of the n fastest content serving sites and m other content serving sites;

receiving data from said fixed location as to the transit times of each of the

n fastest content serving sites and m other content serving sites; and

storing said data in said table.

5   2.   The method of claim 1, wherein said setup phase includes:

receiving a DNS request from said fixed location;

determining a response time for each of the content serving sites;

sending queries to each of the content serving sites, timing said queries so that

they arrive at each of the content serving sites at the same time by using said response

10   time for each of the content serving sites;

receiving data from said fixed location as to the transit times of each of the content

serving sites;

storing said data in a table;

15   3.   The method of claim 2, wherein said determining the response time for each of the

content serving sites includes:

pinging each of the content serving sites;

receiving a reply from each of the content serving sites; and

determining the response time for each of the content serving sites by computing

20   the difference in time from said pinging to said receiving for each of the content serving

sites.

4.   The method of claim 1, wherein said determining the response time for each of n

fastest content serving sites and m other content serving sites includes:

pinging each of n fastest content serving sites and m other content serving sites;

receiving a reply from each of n fastest content serving sites and m other content

5   serving sites; and

determining the response time for each of n fastest content serving sites and m

other content serving sites by computing the difference in time from said pinging to said

receiving for each of the each of said n fastest content serving sites and m other content

serving sites.

10

5.   The method of claim 1, wherein transit time is the time it takes a content serving

site to send a packet to said original location after receiving a query.

6.   The method of claim 1, wherein said data includes at least one record, said record

15   including a transit time, a network address for said fixed location, and a network address

for a content serving site, said transit time being the transit time between said fixed

location and said content serving site.

7.   The method of claim 2, wherein said sending queries to each of the content

20   serving sites includes:

querying each of the content serving sites in order of their response times, longest

response time first, setting a delay for a query for a particular content serving site from

the time the query to the content serving site with the longest response time as one half of said particular content serving site's response time.

8.   The method of claim 1, wherein said sending queries to each of the n fastest

5   content serving sites and m other content serving sites includes:

querying each of said content serving sites n fastest content serving sites and m other content serving sites in order of their response times, longest response time first, setting a delay for a query for a particular content serving site from the time the query to said n fastest content serving site or m other content serving site with the longest response

10   time as one half of said particular content serving site's response time.

9.   The method of claim 1, wherein n is 3 and m is 2.

10.   The method of claim 1, wherein n may be varied based on stability of the

15   computer network.

11.   The method of claim 1, wherein n may be varied based on a rate of change of said content serving sites making up n content serving sites.

20   12.   The method of claim 1, wherein said setup phase occurs when the computer network is first set up.

13.   The method of claim 1, wherein said execution phase occurs when the computer

network is operating normally.

14.   An apparatus for using a phased learning approach for determining the closest

5   content serving sites to a fixed location in a computer network including:

a DNS request receiver coupled to the fixed location;

a response time determiner coupled to said DNS request receiver and coupled to

the content serving sites;

a query sender coupled to said response time determiner and coupled to the

10   content serving sites;

a data receiver coupled to the fixed location;

a memory coupled to said query sender; and

a data storer coupled to said data receiver and coupled to said memory.

15   15.   The apparatus of claim 14, wherein said response time determiner includes an n

fastest content serving site chooser and an m other content serving site chooser.

16.   The apparatus of claim 14, wherein said query sender includes a query timer.

20   17.   The apparatus of claim 14, wherein said response time determiner includes a

pinger and a reply receiver.

18.   The apparatus of claim 14, wherein said query sender includes a response time orderer.

19.   The apparatus of claim 14, further including a phase determiner coupled to said

5   response time determiner, said query sender, and said data receiver.

20.   A program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine for using a phased learning approach for determining the closest content serving sites to a fixed location in a computer network,

10   the method including:

in a setup phase:

storing data as to the transit times of each of the content serving sites in a table;

in an execution phase:

15   receiving a DNS request from said fixed location;

determining a response time for each of n fastest content serving sites and m other content serving sites, said n fastest content serving sites chosen by using said data and choosing the n content serving sites with the lowest transit times, said m other content serving sites chosen by selecting new content serving sites as

20   well as randomly selecting old content serving sites;

sending queries to each of the n fastest content serving sites and m other content serving sites, timing said queries so that they arrive at each of the n fastest

26

content serving sites and m other content serving sites at the same time by using

said response time for each of the n fastest content serving sites and m other

content serving sites;

receiving data from said fixed location as to the transit times of each of the

5   n fastest content serving sites and m other content serving sites; and

storing said data in said table.

21.   An apparatus for using a phased learning approach for determining the closest

content serving sites to a fixed location in a computer network including:

10   means for storing data as to the transit times of each of the content serving sites in

a table in a setup phase;

means for receiving a DNS request from said fixed location in an execution phase;

means for determining a response time for each of n fastest content serving sites

and m other content serving sites, said n fastest content serving sites chosen by using said

15   data and choosing the n content serving sites with the lowest transit times, said m other

content serving sites chosen by selecting new content serving sites as well as randomly

selecting old content serving sites in an execution phase;

means for sending queries to each of the n fastest content serving sites and m other

content serving sites, timing said queries so that they arrive at each of the n fastest

20   content serving sites and m other content serving sites at the same time by using said

response time for each of the n fastest content serving sites and m other content serving

sites in an execution phase;

means for receiving data from said fixed location as to the transit times of each of the *n* fastest content serving sites and *m* other content serving sites in an execution phase; and

means for storing said data in said table in an execution phase.

5

22.   The apparatus of claim 21, further including:

means for receiving a DNS request from said fixed location;

means for determining a response time for each of the content serving sites;

means for sending queries to each of the content serving sites, timing said queries so that they arrive at each of the content serving sites at the same time by using said response time for each of the content serving sites;

means for receiving data from said fixed location as to the transit times of each of the content serving sites;

storing said data in a table;

15

23.   The apparatus of claim 22, wherein said means for determining the response time for each of the content serving sites includes:

means for pinging each of the content serving sites;

means for receiving a reply from each of the content serving sites; and

20   means for determining the response time for each of the content serving sites by computing the difference in time from said pinging to said receiving for each of the content serving sites.

24.   The apparatus of claim 21, wherein said means for determining the response time

for each of n fastest content serving sites and m other content serving sites includes:

means for pinging each of n fastest content serving sites and m other content

5   serving sites;

means for receiving a reply from each of n fastest content serving sites and m

other content serving sites; and

means for determining the response time for each of n fastest content serving sites

and m other content serving sites by computing the difference in time from said pinging

10   to said receiving for each of the each of said n fastest content serving sites and m other

content serving sites.

25.   The apparatus of claim 21, wherein transit time is the time it takes a content

serving site to send a packet to said original location after receiving a query.

15

26.   The apparatus of claim 21, wherein said data includes at least one record, said

record including a transit time, a network address for said fixed location, and a network

address for a content serving site, said transit time being the transit time between said

fixed location and said content serving site.

20

27.   The apparatus of claim 22, wherein said means for sending queries to each of the

content serving sites includes:

29

means for querying each of the content serving sites in order of their response times, longest response time first, setting a delay for a query for a particular content serving site from the time the query to the content serving site with the longest response time as one half of said particular content serving site's response time.

5

28.   The apparatus of claim 21, wherein said means for sending queries to each of the n fastest content serving sites and m other content serving sites includes:

means for querying each of said content serving sites n fastest content serving sites and m other content serving sites in order of their response times, longest response time first, setting a delay for a query for a particular content serving site from the time the query to said n fastest content serving site or m other content serving site with the longest response time as one half of said particular content serving site's response time.

10

29.   The apparatus of claim 21, wherein n is 3 and m is 2.

15

30.   The apparatus of claim 21, wherein n may be varied based on stability of the computer network.

31.   The apparatus of claim 21, wherein n may be varied based on a rate of change of

20   said content serving sites making up n content serving sites.

32.   The apparatus of claim 21, wherein said setup phase occurs when the computer network is first set up.

33.   The apparatus of claim 21, wherein said execution phase occurs when the

5   computer network is operating normally.